Quantitative trait loci for glucosinolate accumulation in *Brassica rapa* leaves

Ping Lou[1], Jianjun Zhao[1,2,3], Hongju He[4,5], Corrie Hanhart[3], Dunia Pino Del Carpio[1], Ruud Verkerk[4], Jan Custers[6], Maarten Koornneef[3,7] and Guusje Bonnema[1]

[1]Laboratory of Plant Breeding, Wageningen University, 6700AJ, Wageningen, the Netherlands; [2]Horticultural College, Hebei Agricultural University, 071001, Baoding, China; [3]Laboratory of Genetics, Wageningen University, 6703BD, Wageningen, the Netherlands; [4]Product Design and Quality Management Group, Wageningen University, 6700EV, Wageningen, the Netherlands; [5]National Engineering Research Center for Vegetables, 100081, Beijing, China; [6]Plant Research International, Wageningen University and Research Centre, 6708PB, Wageningen, the Netherlands; [7]Max Planck Institute for Plant Breeding Research, 50829, Cologne, Germany

Author for correspondence:
Guusje Bonnema
Tel: +31 317 482838
Fax: +31 317 483457
Email: Guusje.Bonnema@wur.nl

Summary

• Glucosinolates and their breakdown products have been recognized for their effects on plant defense, human health, flavor and taste of cruciferous vegetables. Despite this importance, little is known about the regulation of the biosynthesis and degradation in *Brassica rapa*.

• Here, the identification of quantitative trait loci (QTL) for glucosinolate accumulation in *B. rapa* leaves in two novel segregating double haploid (DH) populations is reported: DH38, derived from a cross between yellow sarson R500 and pak choi variety HK Naibaicai; and DH30, from a cross between yellow sarson R500 and Kairyou Hakata, a Japanese vegetable turnip variety.

• An integrated map of 1068 cM with 10 linkage groups, assigned to the international agreed nomenclature, is developed based on the two individual DH maps with the common parent using amplified fragment length polymorphism (AFLP) and single sequence repeat (SSR) markers. Eight different glucosinolate compounds were detected in parents and F_1s of the DH populations and found to segregate quantitatively in the DH populations. QTL analysis identified 16 loci controlling aliphatic glucosinolate accumulation, three loci controlling total indolic glucosinolate concentration and three loci regulating aromatic glucosinolate concentrations.

• Both comparative genomic analyses based on Arabidopsis–*Brassica rapa* synteny and mapping of candidate orthologous genes in *B. rapa* allowed the selection of genes involved in the glucosinolate biosynthesis pathway that may account for the identified QTL.

Introduction

The Brassicaceae form a large plant family with over 3000 species. Six *Brassica* species are cultivated worldwide: three diploids, *Brassica rapa* (AA, $2n = 20$), *B. nigra* (BB, $2n = 16$) and *B. oleracea* (CC, $2n = 18$); and three amphidiploids, *B. juncea* (AABB, $2n = 36$), *B. napus* (AACC, $2n = 38$) and *B. carinata* (BBCC, $2n = 34$). *B. oleracea* and *B. rapa* are among the most important vegetable crops worldwide.

The best characterized phytonutrient compounds in the Brassicaceae are the glucosinolates (Fahey *et al.*, 2001; Wittstock & Halkier, 2002). These secondary metabolites are

predominantly found in plants of the order Brassicales/ Capparales and are normally produced as part of the natural defense against generalist herbivores (Blau *et al.*, 1978; Kliebenstein *et al.*, 2002). As such they generally also have an anti-nutritional value for animals (Mithen, 2001) and great efforts have been made to breed for 'double zero' rapeseed, which is low in erucic acid and the anti-nutritional 3-hydroxy-3-butenyl glucosinolate (Rosa *et al.*, 1997). Remarkably, though, several glucosinolate hydrolysis products, such as isothiocyanates, which are formed upon tissue damage, have an anti-cancer effect when provided in the human diet (Zhang *et al.*, 1994; Fahey *et al.*, 1998; Mithen *et al.*, 2000; Talalay & Fahey, 2001).

The isothiocyanate sulphoraphane, the myrosinase converted form of 4-(methylsulfinyl) butyl glucosinolate (glucoraphanin) present in broccoli, was reported to induce detoxifying enzymes in mammalian cells and acts as a long-lasting antioxidant that detoxifies carcinogens in the body (Zhang *et al.*, 1994; Talalay & Fahey, 2001). Likewise, phenylethyl isothiocyanate, the myrosinase converted form of the main glucosinolate in water-cress, inhibits the action of lung carcinogens (Neo *et al.*, 2005).

The *Brassica* species accumulate aliphatic glucosinolates derived from the amino acid methionine, indole glucosinolates derived from tryptophane and aromatic glucosinolates derived from phenylalanine. In *B. napus*, approx. 30 glucosinolates, and in *B. oleracea* 12 different glucosinolates, have been identified (Mithen *et al.*, 2000; Fahey *et al.*, 2001; Branca *et al.*, 2002). In 58 accessions of *B. rapa* representing all different morphotypes (Chinese cabbage, pak choi, turnips, oil-types and Japanese leafy types like mizuna and komatsuna), eight different glucosinolates were detected (He *et al.*, 2000, unpublished), while in 113 different accessions of turnip greens (*Brassica rapa* subsp. *rapa* L.), 16 different glucosinolates were detected (Padilla *et al.*, 2007). Despite the wide variation in glucosinolate profiles among genotypes, individual plants mostly contain a limited number of major glucosinolate forms. In all *B. rapa* genotypes studied, the aliphatic (4-C) 3-butenyl and the aliphatic (5-C) 4-pentenyl glucosinolates were the predominant glucosinolates.

In the species *A. thaliana*, more than 34 different glucosinolates have been identified (Kliebenstein *et al.*, 2001a; Reichelt *et al.*, 2002). The glucosinolate biosynthetic pathway consists of three stages in Arabidopsis and *Brassica* ssp., namely an amino-acid chain elongation stage for methionine- and phenylalanine-derived glucosinolates, a glucosinolate skeleton formation and a side-chain modification stage (Halkier & Du, 1997; Mithen, 2001; Halkier & Gershenzon, 2006). Upon tissue disruption, glucosinolates come into contact with the plant enzyme myrosinase, resulting in the formation of many, often volatile, glucosinolate degradation products with diverse biological activities.

To date, the structural genes responsible for most biosynthetic steps have been identified and the core pathway of glucosinolate biosynthesis has been clarified in Arabidopsis (Wittstock & Halkier, 2000; Bak & Feyereisen, 2001; Grubb *et al.*, 2004; Mikkelsen *et al.*, 2004; Piotrowski *et al.*, 2004; Halkier & Gershenzon, 2006). Genetic studies in Arabidopsis and *Brassica* spp. proved an important role for AOP (2-oxoglutarate-dependent dioxygenase) and MAM (methyl-thioalkylmalate synthase) in the regulation of aliphatic glucosinolate accumulation. The *MAM* gene family is involved in glucosinolate side-chain elongation, and the *AOP* family is involved in side-chain modification (Kliebenstein *et al.*, 2001a,b; Kroymann *et al.*, 2001, 2003; Field *et al.*, 2004, 2006; Heidel *et al.*, 2006).

In *B. oleracea*, glucosinolate-Elong (MAM) and glucosinolate-Alk (AOP), responsible for glucosinolate side-chain elongation and modification in Arabidopsis, were cloned based on Arabidopsis homology (Li & Quiros, 2002; 2003) and the organization of AOP genes was further compared with Arabidopsis sequences (Gao *et al.*, 2004). Additionally, *BoGSL-Alk* expressed in Arabidopsis resulted in changed aliphatic glucosinolate profiles in leaves and seeds (Li & Quiros, 2003).

Besides these and other structural genes, several transcription factors were identified as regulators of indole glucosinolate or aliphatic glucosinolate accumulation, such as IQD1, ART1/Myb34, AtDof1.1(OBP2), HAG1/Myb28, HAG3/Myb29, HAG2/Myb76 and HIG1/Myb51 (Celenza *et al.*, 2005; Hirai *et al.*, 2005, 2007; Levy *et al.*, 2005; Skirycz *et al.*, 2006; Gigolashvili *et al.*, 2007a,b,c; Sonderby *et al.*, 2007; Wentzell *et al.*, 2007).

Genetic factors strongly influence glucosinolate accumulation, although environmental and developmental factors may modify the expression of genes. Quantitative trait locus (QTL) analysis is a powerful method to study the genetics underpinning quantitative variation in glucosinolate profiles, and to estimate the number of variable loci affecting a trait (Koornneef *et al.*, 2004). In Arabidopsis, one QTL analysis was performed for glucosinolate accumulation in both seeds and leaves using L*er* × Cvi recombinant inbred lines. Total aliphatic glucosinolate accumulation was controlled by six QTL, total indolic glucosinolate accumulation was controlled by another six QTL and benzylic glucosinolate accumulation was controlled by three QTL (Kliebenstein *et al.*, 2001c). Kliebenstein *et al.* (2001c) showed that two QTL controlling total aliphatic glucosinolates that map to *AOP* and GS-elong (*MAM*) loci interact epistatically, which was also observed in expression QTL analysis by Keurentjes *et al.* (2006) and Wentzell *et al.* (2007).

In the L*er* × Col RI population there was much less variation for total aliphatic glucosinolate accumulation and only two chromosome 5 QTL were identified, with the major QTL explaining 60% of the variation and colocalizing with *MAM* (Kliebenstein *et al.*, 2002). QTL for glucosinolate accumulation were also identified in a new Arabidopsis inbred line population (Da(1)-12 × Ei-2 (Pfalz *et al.*, 2007). Total aliphatic glucosinolates in this population were controlled by three QTL, one co-localizing with AOP, one near the top of chromosome 5, which was also identified previously, and a new QTL on chromosome 1. Genetic analysis of a series of Arabidopsis

ecotypes revealed the *MAM* locus at the top of chromosome 5, explaining a large proportion of the observed variation in aliphatic glucosinolate composition (Kroymann *et al.*, 2003).

In *Arabidopsis lyrata* interpopulation crosses, one QTL was identified at the *MAM* locus for aliphatic glucosinolate ratios plus an additional QTL for total indole glucosinolates and the ratio of aliphatic to indole glucosinolates (Heidel *et al.*, 2006). Multiple QTL have been described both for seed and leaf glucosinolates in both *B. napus* (Howell *et al.*, 2003) and *B. juncea* (Mahmood *et al.*, 2003; Ramchiary *et al.*, 2007).

In *B. rapa*, there has been no report on QTL controlling glucosinolate accumulation until now. The objectives of this study are to map QTL for glucosinolate composition and accumulation in *B. rapa* leaves in two new double haploid (DH) populations derived from wide crosses.

Materials and methods

Parental materials used to create DH populations

Brassica rapa accession yellow sarson YS-143 (accession number FIL500) was crossed as male parent to both pak choi PC-175 (cv. Nai Bai Cai; accession number VO2B0226) and vegetable turnip VT-115 (cv. Kairyou Hakata; accession number CGN15199) to develop DH populations DH38 and DH30, respectively. These accessions were selected based on their different morphotypes, their genetic distances as determined by AFLP fingerprinting and their phytate contents (Zhao *et al.*, 2005, 2007; Lou *et al.*, 2007).

Developing DH populations using microspore culture

The microspore culture protocol was based on Coventry *et al.* (1988) and Custers *et al.* (1994, 2001). The donor plants were grown in the glasshouse and transferred to a growth cabinet at a 15 : 12°C day : night temperature regime at the beginning of bolting. After the first one to three flowers had opened, inflorescences were collected for microspore isolation. Buds were excised and were surface-sterilized by a short dip in 70% ethanol and, thereafter, treated in 2% NaOCl for 10 min, and they were then rinsed three times with sterile distilled water for *c.* 1, 4 and 10 min each. The buds were gently macerated in a tube containing NLN13 liquid medium, and microspores were collected by filtration through two layers of 44 μm nylon filter and then washed three times by centrifugation at 100 *g* for 3–4 min each time. After the last washing, the microspores were resuspended at a density of 4×10^4 ml^{-1} in NLN13 medium and the suspension was poured into Petri dishes for culturing. Cultures were incubated at 31°C for 72 h in complete darkness, and then transferred to 25°C, again in the dark. Finally two to eight embryos from each Petri dish were germinated and rooted on B5-Gamborg medium with 1% sucrose, and well-rooted plants were transferred to the glasshouse.

Spontaneously diploidized plants were used for the further experiments. Flower stalks were bagged to prevent cross-pollination, and bud pollination was applied in a number of cases for self-incompatible plants that did not set seed spontaneously. The progeny of the DH plants from one single F_1 plant per crossing combination were used for genotyping and phenotyping.

Two DH populations were used to construct linkage maps: DH38 consisted of 71 lines and DH30 consisted of 64 lines and they shared one common male parent YS-143.

Genotyping

Leaf material for DNA extraction was collected from seedlings and then ground using a Retsch 300 shaker (Retsch BV, Ochten, the Netherlands). DNA was isolated based on modified CTAB methods (Murray & Thompson, 1980).

The AFLP analysis was performed according to Vos *et al.* (1995) using the enzyme combinations *Eco*RI/*Mse*I and *Pst*I/*Mse*I. The primer combinations used for the mapping were those combinations used to construct the phenetic tree (Zhao *et al.*, 2005) plus a selection based on a pre-screening of the parental accessions. In total, 15 *Eco*RI + 3/*Mse*I + 2 and *Pst*I + 2/*Mse*I + 3 primer combinations were used. The primer combinations were named according to the standard primer combination code. AFLP fingerprints were visualized on a NEN Global Edition IR^2 DNA Analyser (Li-Cor Bioscience, Lincoln, NE, USA) by using IRDye 700 and IRDye 800 labelled *Eco*RI + 3 or *Pst*I + 2 primers. AFLP marker nomenclature is based on the enzyme combination (E = *Eco*RI, P = *Pst*I and M = *Mse*I) followed by the size of the band. Yellow sarson markers were followed by a Y, while markers that were absent in YS, but had identical band sizes in PC-175 and VT-115, were followed by a y.

Sequence information of a collection of 93 microsatellites was obtained from the Multinational Brassica Genome Project (MBGP) (www.brassica.info) website. The SSRs are derived from different genomic sources (Arabidopsis, *B. rapa*, *B. oleracea*, *B. napus* and *B. nigra*, respectively, prefixed by Nga, Ra/BRMS, Ol/sORA, Na/Bn and Ni). Three SSR primers (BRH80A08FLC1, BrH80A08FLC2, BRH80C09MAF) were kindly provided by Dr Jungsun Kim (Kim *et al.*, 2006), and are developed from BACs from the KBrH library (*Hind*III library of *Brassica rapa* inbred line Chiifu) containing *BrFLC* sequences (Yang *et al.*, 2006). Two markers were developed from BACs containing AOP loci; one SCAR marker (SCAR-AOPc) with primers ODD12/ODD48 was based on the *AOP* gene on a *B. oleracea* BAC (B21H13, GenBank accession number AC122543) (Gao *et al.*, 2004) and one SSR (SSR-AOPa) based on a *B. rapa* BAC harbouring the AOP locus (KBrB002P01, *Bam*HI library of *B. rapa* inbred line Chiifu) (SSR-AOPa: forward: ATGTGTTTATTGTAACCGGG, reverse: GCGTAAACATGATGGAATTT; Dr Jung Sun Kim, pers. comm.).

Table 1 Traits, nomenclature[a] and abbreviation

Abbreviation	Description or common name	Chemical names or explanation
Aliphatic glucosinolates		
PRO	Progoitrin	2-hydroxy-3-butenyl glucosinolate (4 C)
NAP	Gluconapin	3-butenyl glucosinolate (4-C)
4-C ali	PRO+NAP	
GBN	Glucobrassicanapin	4-pentenyl glucosinolate (5-C)
5-C ali	GBN	
Tali	Total aliphatic glucosinolates	The sum of PRO, NAP and GBN
Ratio1	(PRO+NAP)/Tali	Ratio for side-chain elongation
Ratio2	PRO/(PRO+NAP)	Ratio for side-chain modification
Indole glucosinolates		
4ME	4-Methoxy glucobrassicin	4-Methoxyindol-3-ylmethyl glucosinolate
4OH	4-Hydroxy glucobrassicin	4-Hydroxyindol-3-ylmethyl glucosinolate
NEO	Neoglucobrassicin	1-Methoxyindol-3-ylmethyl glucosinolate
GBC	Glucobrassicin	Indol-3-ylmethyl glucosinolates
Tind	Total indolic glucosinolates	The sum of 4ME,4OH,NEO and GBC
Aromatic glucosinolates		
NAS	Gluconasturtiin	2-Phenylethyl glucosinolate
Aro	Total aromatic glucosinolates	Equal to NAS in this experiment

[a]Nomenclature based on Talalay & Fahey (2001).

Map construction

Linkage analysis and map construction werecarried out using JoinMap 4.0 (http://www.kyazma.nl) for each population. Linked loci were grouped on the basis of pairwise LOD values between 4 and 7. The Kosambi mapping function was used to convert recombination data to map distances. Locus order within LOD grouping was decided based on RECORD and JoinMap4.0 as described by Van Os *et al.* (2005) and Van Ooijen (2006), respectively. Linkage maps of the two mapping populations were integrated according to the principle described by Stam (1993) using Haldane's mapping function. The linkage maps were drawn using MapChart 2.1(Voorrips, 2002). Linkage groups were assigned based on SSR markers that are also mapped in reference maps. The 10 linkage groups of *B. rapa* were renamed A1–A10, corresponding to R01–R10, respectively, as decided at the Multinational Steering Committee January 2007. The map orientation presented by Choi *et al.*, (2007) was taken as reference map (www.brassica.info).

Glucosinolate isolation and determination

The two DH populations were grown in autumn 2004 (24 September–15 November) and spring 2005 (3 March–11 April) in a heated glasshouse in Wageningen, the Netherlands, under controlled day length of 16 h and 24 : 18°C day : night temperature. Since there was no cooling in the glasshouse, during the daytime temperatures were occasionally higher than 24°C. In the autumn 2004 experiment, one fully expanded leaf from each pair of plants grown side by side per DH line was harvested *c.* 50 d after sowing, frozen in liquid nitrogen and stored at −80°C. In the spring of 2005, fully expanded leaves were harvested from three plants grown side by side per DH line 40 d after sowing and stored at −26°C. Leaf material per DH line per experiment was combined, freeze-dried using the freeze-drying instrument GRI 20-85 MP and ground into a fine powder to represent one biological replication per experiment. The freeze-dried plant material was stored at −30°C for further glucosinolate analysis.

Sample extraction, desulfation and HPLC analysis were performed according to Verkerk *et al.* (2001). The results were analyzed using the ChromQuest program (Thermo Electron Corporation, Middelburg, the Netherlands). The glucosinolate content of each sample was calculated using the response factors referring to the internal standard (benzyl glucosinolate). The results were expressed as μmol g^{-1} dry weight (DW) of material. The peaks were identified by comparison with individual glucosinolate standards and with data obtained from the literature.

QTL analysis

The software program MAPQTL 5.0 (Van Ooijen, 2004) was used to detect QTL with both the interval mapping and MQM mapping method on DH30 and DH38. In the regions of the putative QTL, the markers with the highest LOD values were taken as co-factors. The LOD thresholds for QTL significance were determined by permutation tests (1000 replications) in MapQTL with a genome-wide significance level of $P = 0.05$ for significant linkages; LOD = 2.0 was used as significant threshold for both DH38 and DH30 based on the permutation test.

For QTL analysis, all the traits, ratios and abbreviations are listed in Table 1. The ratio between aliphatic glucosinolates of

different side-chain lengths was calculated as (PRO + NAP)/Tali. The ratio between the side-chain modification from 3-butenyl (4-C NAP) to 2-hydroxy-3-butenyl (4-C PRO) glucosinolates was calculated as PRO/(NAP + PRO).

Quantitative trait loci are graphically displayed, including 1-LOD and 2-LOD support intervals, using mapchart2.1 (Voorrips, 2002). Frequency distributions and correlation coefficients were obtained and ANOVA analysis was performed using Genstat 8.1. Nomenclature for QTL was as follows: name for glucosinolate type (Ali, Ind or Aro), followed by linkage group number and relative position of the QTL from the top to the bottom of that linkage group.

Results

Development of DH populations and construction of linkage maps

Genetic linkage maps were constructed for two DH populations having YS-143 as the common male parent and PC-175 (DH38) or VT-115 (DH30) as the female parent, which resulted in hybrids that responded well in microspore culture. A total of 321 AFLP and SSR loci were mapped in DH38, and 300 AFLP and SSR loci were mapped in DH30. The total map lengths were 688 and 763 cM, with average distances between markers of 2.14 and 2.30 cM, respectively. SSR primers were selected from the public databases to represent each linkage group at least once, but preferably twice. From a total of 96 mapped SSR primers screened against the three parental lines, 33 displayed polymorphisms between the two pairs of parent lines, 22 of which were mapped. Each of the 10 linkage groups has at least one SSR marker, which allows anchoring to existing published maps, and in case more SSR markers were mapped within linkage groups, the orientation could be determined compared with a published *B. rapa* map (Kim *et al.*, 2006; Choi *et al.*, 2007). Additionally we mapped two markers developed from BACs that encoded AOP genes, one SSR (SSR-AOPa) located on A3 and one SCAR (SCAR-AOPc) located on the top of A9.

The integration of linkage groups and assignments to linkage groups of the two DH maps was based on 116 AFLP markers from the common yellow sarson parent and 14 SSR markers. To compare linkage group orientation, we used the reference map published by Choi *et al.* (2007). Since linkage groups A1, A4 and A8 only contain one SSR, their orientation compared to this reference map cannot be inferred. Map integration is useful to compare the QTL position between the two different populations. The total number of loci in the integrated map was 470 and the map covers 1068 cM (Haldane distance) of the *B. rapa* genome, divided over 10 linkage groups. The average interval between markers was 2.27 cM (Fig. 1). Two gaps over 20 cM were present in A5 and A10. A9 contained a large cluster of linked markers and was the longest linkage group with 159.5 cM. A1 and A4 were

the shortest linkage groups (63.2 and 72.8 cM, respectively). Significant segregation distortion from the expected Mendelian segregation ratio (1 : 1) ($P < 0.01$, using a chi square test) were observed for 30–40% of all markers in both DH populations, in which the marker loci exhibiting distorted segregation were not randomly distributed throughout the genome but were clustered on several linkage groups (Fig. 1).

Quantitative variation in glucosinolates content

We conducted two independent experiments, one in autumn 2004 using leaves of 50-d-old plants, and one in spring 2005 using leaves of 40-d-old plants. HPLC results showed that the three parental accessions, their F_1s and lines of the two DH populations accumulated a set of eight glucosinolates. The eight different glucosinolates identified in these genotypes belong to three classes: the aliphatic glucosinolates, the indolic glucosinolates and one aromatic glucosinolate. The aliphatic glucosinolate 3-butenyl glucosinolate (NAP) was the most dominant glucosinolate in *B. rapa*, with maximum concentrations around 51.0 μmol g^{-1} DW in DH-30 in autumn 2004, followed by 4-pentenyl glucosinolate (GBN) (8.6 μmol g^{-1} DW in VT-115 in autumn 2004) and 2-hydroxy-3-butenyl glucosinolate (PRO) (2.3 μmol g^{-1} DW in DH-30, spring 2005). The aromatic glucosinolate 2-phenylethyl glucosinolate (NAS) had concentrations between 0.1 and 3.0 μmol g^{-1} DW. The indolic glucosinolates, 3-indolylmethyl (GBC), 4-methoxy-3-indolylmethyl (4ME), 4-hydroxy-3-indolylmethyl (4OH) and 1-methoxy-3-indolylmethyl (NEO) glucosinolate, were present in lower quantities and are represented as total indole glucosinolates (Table 2).

The DH lines of both populations showed a continuous distribution for glucosinolate accumulation, suggesting multigenic control of glucosinolates in *B. rapa* leaves (Fig. 2). In the case of DH38, the variation among the DH lines for total aliphatic glucosinolates is between the parental values, whereas in DH30, transgression to lower values is observed. Indolic glucosinolates show strong transgression to higher values in DH38, but this is less pronounced in DH30, even though the highest values are found in DH30. The yellow sarson parent YS-143 has intermediate aliphatic and indole glucosinolates concentrations compared with the pak choi parent (lower) and the vegetable turnip parent (higher). In many cases, the F_1 hybrids, which were only analyzed in spring 2005, show the highest values, indicating dominance of high glucosinolate concentrations.

Correlations were calculated between the two experiments and between individual glucosinolates within experiments and were always positive, except between NAP and PRO, and between NAP and NAS in 2005 (Table 3 and Supporting Information, Table S1). The correlation between 2004 and 2005 was significant at $P < 0.001$ and $P < 0.0001$ for the aliphatic glucosinolates and for the indole glucosinolates GBC and 4ME, while the 2004–2005 correlation was only significant at $P < 0.01$ for the aromatic GS NAS and the indole glucosinolate
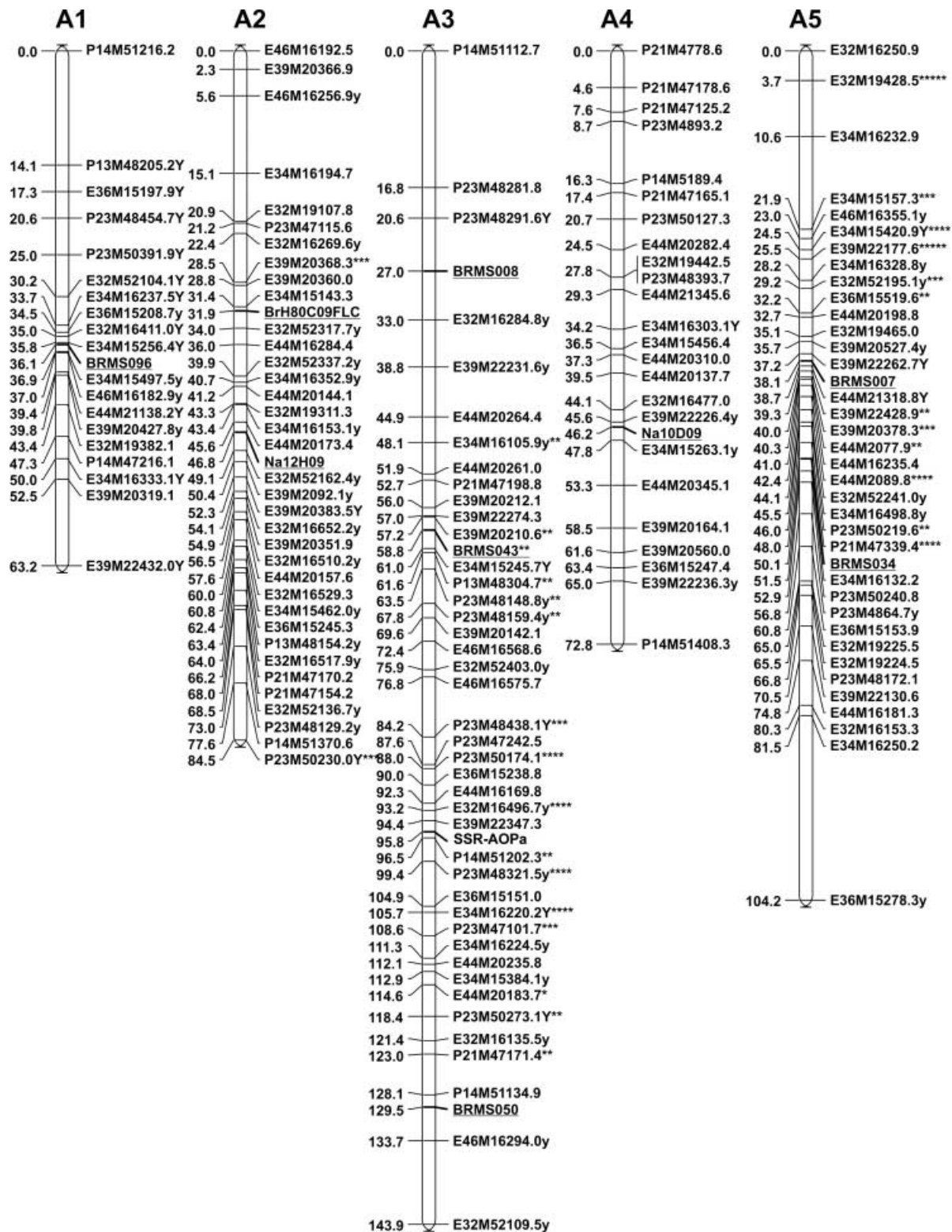
Fig. 1 Integrated linkage map of *Brassica rapa* based on the DH38 and DH30 populations with amplified fragment length polymorphism (AFLP) and anchor single sequence repeat (SSR) markers. Anchor SSR markers are underlined. Linkage groups are numbered according to international standard A1–A10. The degree of segregation distortion (**, *P* < 0.01; ***, *P* < 0.001; ****, *P* < 0.0001; *****, *P* < 0.00001) is represented by an asterisk next to marker locus names. AFLP marker names first list the enzyme/primer combination (E**M** or P**M**, followed by the size of the polymorphic fragment. Next to the size, a capital Y refers to a yellow sarson marker segregating in both DH38 and DH30, while a small y refers to a segregating marker with identical size in PC and VT, and absent in yellow sarson.
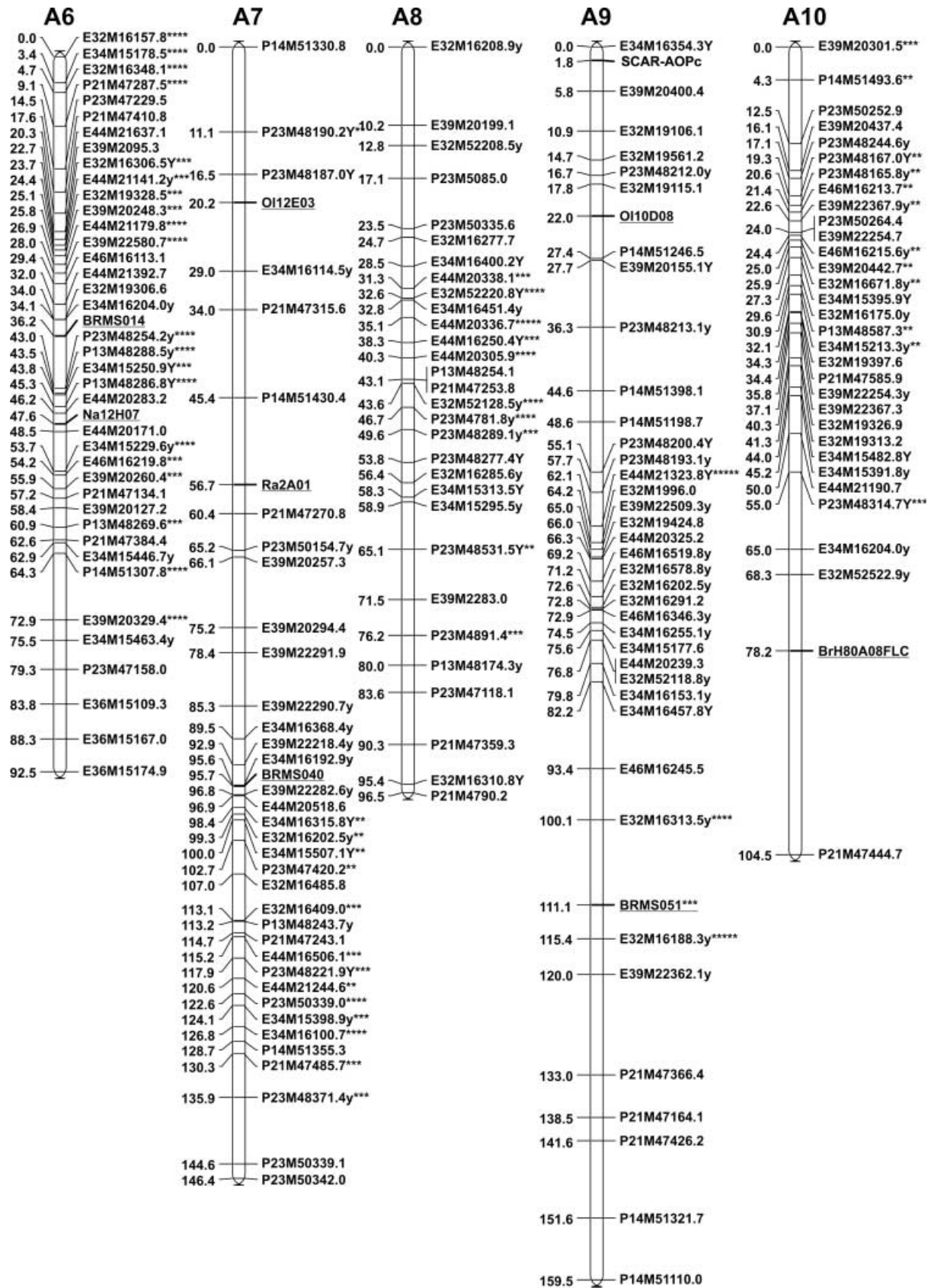
A6

Position	Marker
0.0	E32M16157.8****
3.4	E34M15178.5****
4.7	E32M16348.1****
9.1	P21M47287.5****
14.5	P23M47229.5
17.6	P21M47410.8
20.3	E44M21637.1
22.7	E39M2095.3
23.7	E32M16306.5Y***
24.4	E44M21141.2y***
25.1	E32M19328.5***
25.8	E39M20248.3***
26.9	E44M21179.8****
28.0	E39M22580.7****
29.4	E46M16113.1
32.0	E44M21392.7
34.0	E32M19306.6
34.1	E34M16204.0y
36.2	BRMS014
43.0	P23M48254.2y****
43.5	P13M48288.5y****
43.8	E34M15250.9Y***
45.3	P13M48286.8Y****
46.2	E44M20283.2
47.6	Na12H07
48.5	E44M20171.0
53.7	E34M15229.6y****
54.2	E46M16219.8***
55.9	E39M20260.4****
57.2	P21M47134.1
58.4	E39M20127.2
60.9	P13M48269.6***
62.6	P21M47384.4
62.9	E34M15446.7y
64.3	P14M51307.8****
72.9	E39M20329.4****
75.5	E34M15463.4y
79.3	P23M47158.0
83.8	E36M15109.3
88.3	E36M15167.0
92.5	E36M15174.9

A7

Position	Marker
0.0	P14M51330.8
11.1	P23M48190.2Y*
16.5	P23M48187.0Y
20.2	OI12E03
29.0	E34M16114.5y
34.0	P21M47315.6
45.4	P14M51430.4
56.7	Ra2A01
60.4	P21M47270.8
65.2	P23M50154.7y
66.1	E39M20257.3
75.2	E39M20294.4
78.4	E39M22291.9
85.3	E39M22290.7y
89.5	E34M16368.4y
92.9	E39M22218.4y
95.6	E34M16192.9y
95.7	BRMS040
96.8	E39M22282.6y
96.9	E44M20518.6
98.4	E34M16315.8Y**
99.3	E32M16202.5y**
100.0	E34M15507.1Y**
102.7	P23M47420.2**
107.0	E32M16485.8
113.1	E32M16409.0***
113.2	P13M48243.7y
114.7	P21M47243.1
115.2	E44M16506.1***
117.9	P23M48221.9Y***
120.6	E44M21244.6**
122.6	P23M50339.0****
124.1	E34M15398.9y***
126.8	E34M16100.7****
128.7	P14M51355.3
130.3	P21M47485.7***
135.9	P23M48371.4y***
144.6	P23M50339.1
146.4	P23M50342.0

A8

Position	Marker
0.0	E32M16208.9y
10.2	E39M20199.1
12.8	E32M52208.5y
17.1	P23M5085.0
23.5	P23M50335.6
24.7	E32M16277.7
28.5	E34M16400.2Y
31.3	E44M20338.1***
32.6	E32M52220.8Y****
32.8	E34M16451.4y
35.1	E44M20336.7*****
38.3	E44M16250.4Y***
40.3	E44M20305.9****
43.1	P13M48254.1 / P21M47253.8
43.6	E32M52128.5y****
46.7	P23M4781.8y****
49.6	P23M48289.1y***
53.8	P23M48277.4Y
56.4	E32M16285.6y
58.3	E34M15313.5Y
58.9	E34M15295.5y
65.1	P23M48531.5Y**
71.5	E39M2283.0
76.2	P23M4891.4***
80.0	P13M48174.3y
83.6	P23M47118.1
90.3	P21M47359.3
95.4	E32M16310.8Y
96.5	P21M4790.2

A9

Position	Marker
0.0	E34M16354.3Y
1.8	SCAR-AOPc
5.8	E39M20400.4
10.9	E32M19106.1
14.7	E32M19561.2
16.7	P23M48212.0y
17.8	E32M19115.1
22.0	OI10D08
27.4	P14M51246.5
27.7	E39M20155.1Y
36.3	P23M48213.1y
44.6	P14M51398.1
48.6	P14M51198.7
55.1	P23M48200.4Y
57.7	P23M48193.1y
62.1	E44M21323.8Y*****
64.2	E32M1996.0
65.0	E39M22509.3y
66.0	E32M19424.8
66.3	E44M20325.2
69.2	E46M16519.8y
71.2	E32M16578.8y
72.6	E32M16202.5y
72.8	E32M16291.2
72.9	E46M16346.3y
74.5	E34M16255.1y
75.6	E34M15177.6
76.8	E44M20239.3 / E32M52118.8y
79.8	E34M16153.1y
82.2	E34M16457.8Y
93.4	E46M16245.5
100.1	E32M16313.5y****
111.1	BRMS051***
115.4	E32M16188.3y*****
120.0	E39M22362.1y
133.0	P21M47366.4
138.5	P21M47164.1
141.6	P21M47426.2
151.6	P14M51321.7
159.5	P14M51110.0

A10

Position	Marker
0.0	E39M20301.5***
4.3	P14M51493.6**
12.5	P23M50252.9
16.1	E39M20437.4
17.1	P23M48244.6y
19.3	P23M48167.0Y**
20.6	P23M48165.8y**
21.4	E46M16213.7**
22.6	E39M22367.9y**
23.?	P23M50264.4
	E39M22254.7
24.0	
24.4	E46M16215.6y**
25.0	E39M20442.7**
25.9	E32M16671.8y**
27.3	E34M15395.9Y
29.6	E32M16175.0y
30.9	P13M48587.3**
32.1	E34M15213.3y**
34.3	E32M19397.6
34.4	P21M47585.9
35.8	E39M22254.3y
37.1	E39M22367.3
40.3	E32M19326.9
41.3	E32M19313.2
44.0	E34M15482.8Y
45.2	E34M15391.8y
50.0	E44M21190.7
55.0	P23M48314.7Y***
65.0	E34M16204.0y
68.3	E32M52522.9y
78.2	BrH80A08FLC
104.5	P21M47444.7

Fig. 1 continued

Traits[a]	Parental lines			F$_1$s		DH38			DH30		
	YS-143	PC-175	VT-115	F1-38	F1-30	Avr	Min.	Max.	Avr	Min.	Max.
PRO04	1.3	0.8	0.7	na	na	0.4	0.1	1.9	0.4	0.2	1.0
PRO05	0.8	0.9	0.6	2.2	1.4	0.7	0.2	1.8	0.5	0.2	2.3
NAP04	22.2	0.6	42.1	na	na	3.0	0.1	18.5	20.0	4.0	51.0
NAP05	22.3	0.5	40.1	18.3	47.4	3.4	0.2	17.9	14.0	4.0	37.0
GBN04	0.2	0.3	7.8	na	na	0.7	0.1	5.7	2.1	0.4	8.1
GBN05	0.2	0.3	8.6	4.1	3.0	0.6	0.1	3.4	0.8	0.1	3.3
Tali04	35.2	1.7	50.6	na	na	4.1	0.2	26.1	22.5	4.6	60.2
Tali05	23.0	1.7	48.6	24.2	51.2	4.7	0.4	23.1	15.3	4.4	42.6
Tind04	0.5	0.2	1.9	na	na	0.8	0.1	2.9	1.3	0.4	4.7
Tind05	0.4	0.3	1.5	0.5	1.1	0.6	0.1	2.0	0.9	0.3	2.8
Aro04	0.8	0.4	0.7	na	na	0.8	0.1	2.9	0.6	0.1	3.0
Aro05	0.6	0.5	0.8	2.3	1.0	1.0	0.4	2.4	0.3	0.1	0.6

[a]Trait abbreviations as explained in Table 1, followed by the year; 04, autumn 2004; 05, spring 2005.
na, not applicable.
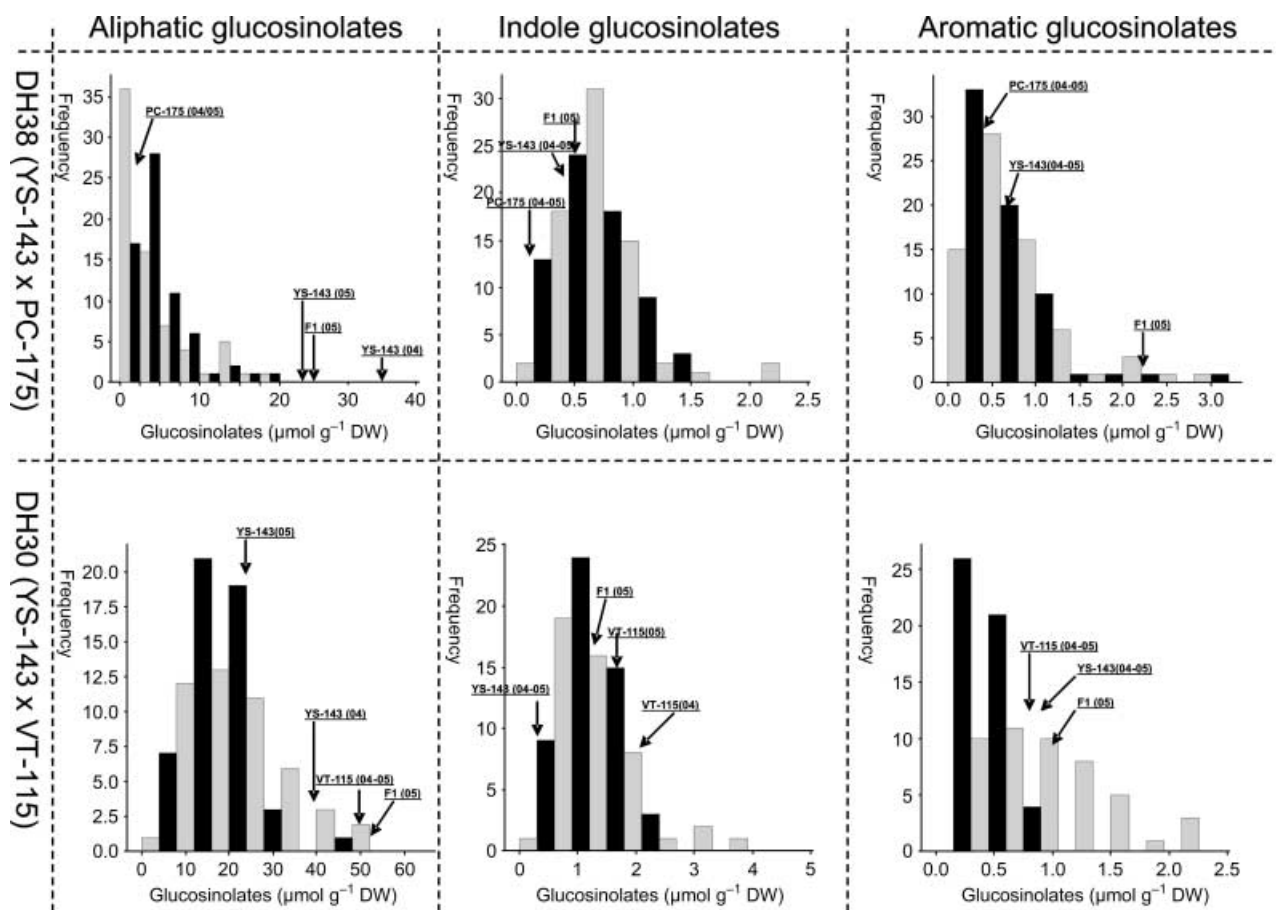
Fig. 2 Frequency distribution of aliphatic, indolic and aromatic glucosinolates in *Brassica rapa* double haploid (DH) populations DH38 (YS-143xPC-175) and DH30 (YS-143xVT-115) in two experiments. Gray bar, autumn 2004; black bar, spring 2005. Parental values and F$_1$ are indicated with arrows with the year.

Table 3 Correlation coefficients of individual glucosinolate concentrations of DH30 and DH38 populations in 2004 and 2005

2005 \ 2004	PRO04	NAP04	GBN04	4OH04	GBC04	4ME04	NEO04	NAS04	2004 \ 2004
PRO05	0.42**	0.07	0.20	0.05	0.18	0.13	0.12	0.04	PRO04
NAP05	−0.02	0.69***	0.55***	0.26*	0.41**	0.21	0.25*	0.22	NAP04
GBN05	0.23	0.22	0.55***	0.28*	0.55***	0.59***	0.25	0.41**	GBN04
4OH05	0.01	0.24	0.30*	0.25*	0.27*	0.20	0.18	0.23	4OH04
GBC05	0.05	0.47**	0.36**	0.29*	0.58***	0.57***	0.41**	0.42**	GBC04
4ME05	0.21	0.47**	0.44**	0.36**	0.74***	0.48**	0.28*	0.54***	4ME04
NEO05	0.11	0.02	0.07	0.33**	0.24	0.30*	0.13	0.42*	NEO04
NAS05	0.16	−0.18	0.12	0.26*	0.14	0.07	0.32**	0.25*	NAS04
2005 \ 2005	PRO05	NAP05	GBN05	4OH05	GBC05	4ME05	NEO05	NAS05	2004 \ 2005

Upper triangle, correlation coefficients of glucosinolates in 2004; lower triangle, correlation coefficients of glucosinolates in 2005. The middle beam shows the correlation of each individual glucosinolate between 2004 and 2005.
*, $P < 0.01$; **, $P < 0.001$; ***, $P < 0.0001$.

4OH, and was not significant for the fourth indole glucosinolate NEO. The concentrations of several individual indole glucosinolates were significantly correlated and a significant correlation was also observed between the aliphatic GBN and several indole glucosinolate concentrations in both experiments (Table 3). Within the different experiments, the aromatic glucosinolate NAS is correlated with three indole glucosinolates and with the aliphatic glucosinolate GBN in autumn 2004 (Table 3). The aliphatic glucosinolate PRO showed no significant correlation to any of the glucosinolates, while the most abundant glucosinolate NAP showed correlation with GBN in 2004 and with several indole glucosinolates in autumn 2004 and spring 2005.

QTL controlling total and individual aliphatic glucosinolate concentrations

To investigate the genetic control of quantitative variation of aliphatic glucosinolates in *B. rapa* leaves, we analyzed total and individual aliphatic glucosinolate concentrations in both DH populations in the two different seasons. A total of 16 QTL for different aliphatic glucosinolates and their ratios were identified on all linkage groups except for A2, with major QTL on A3 and A10 (Table 4, Table S2). These include eight QTL for the predominant glucosinolate NAP, two for GBN and five QTL for the least abundant aliphatic glucosinolate PRO (Table 4). Six QTL were identified for total aliphatic glucosinolate concentrations, four of which also affected NAP (Ali-QTL3.2, −6.1, −7.1 and −10.1). The sum of these QTL explained 40 and 56% of the phenotypic variation of total leaf aliphatic glucosinolate content in DH38 in the different

seasons (autumn 2004 and spring 2005, respectively), and 75 and 47% in DH30 (autumn 2004 and spring 2005, respectively). The major QTL Ali-QTL3.2 on the bottom of A3 explained 42, 26 and 44% of the variation for the aliphatic glucosinolates NAP, GBN and total aliphatic glucosinolates in spring 2005, but only 12% for NAP in autumn 2004, while for GBN and total aliphatic glucosinolates, no significant QTL were identified. In DH30 we see the inverse, since this QTL explained only 25 and 22% of the variation for NAP and GBN in spring 2005, but much higher percentages of 56, 22 and 51% for NAP, GBN and total aliphatic glucosinolates, respectively, in autumn 2004. For Ali-QTL3.2, the YS-143 alleles increase the aliphatic glucosinolate concentration in DH38, while in DH30, the parental VT-115 alleles (and not YS-143) increase aliphatic glucosinolate concentrations. Ali-QTL3.2 also affects ratio 1 between the 4C glucosinolates NAP plus PRO and Tali (4C and 5C) and ratio 2 between PRO and PRO plus NAP (side-chain modification). In the middle of A3, another QTL, Ali-QTL3.1, was identified for aliphatic glucosinolate PRO in DH38, which was not detected at all in DH30. Also in DH38, the YS-143 alleles increase the PRO concentration. Besides Ali-QTL3.2, only Ali-QTL3.1 and Ali-QTL7.1 were identified in both DH populations. Ali-QTL7.1 explained 13 and 32% variation for total aliphatic glucosinolates in the spring 2005 experiment for DH38 and DH30, respectively. All other QTL for aliphatic glucosinolates were identified in only one DH population and often in only one experiment, and generally explained between 6 and 20% of the phenotypic variation for total or individual glucosinolates, except for Ali-QTL10.1, which explained 43% of the variation for PRO in DH38 in autumn 2004 only.

Table 4 Summary of all quantitative trait loci (QTL) detected in two *Brassica rapa* populations in 2004 and 2005

Name[a]	Trait[b]	Position	DH38-04au		DH38-05sp		DH30-04au		DH30-05sp		QTL effect[c]	
			LOD	Exp%	LOD	Exp%	LOD	Exp%	LOD	Exp%	DH38	DH30
Ali-QTL1.1	PRO	A1-Top	N	N	3.4	17.2	N	N	N	N	+	N
Ali-QTL1.2	Tali	A1-Mid	3.0	9.0	N	N	N	N	N	N	−	N
	NAP	A1-Mid	3.0	9.1	*1.0*	*1.9*	N	N	N	N	+	N
Ali-QTL1.3	Ratio2	A1-Bot	2.0	11.5	N	N	N	N	N	N	−	N
Ali-QTL3.1	Ratio1	A3-Mid	4.3	12.1	3.5	6.4	1.3	7.8	2.7	11.7	−	+
	Ratio2	A3-Mid	*1.8*	*11.0*	9.0	28.0	N	N	N	N	−	N
	NAP	A3-Mid	*1.5*	*3.0*	3.0	7.2	N	N	N	N	+	N
	PRO	A3-Mid	2.4	12.5	5.5	30.7	N	N	N	N	−	N
Ali-QTL3.2	NAP	A3-Bot	4.1	12.4	11.4	42.0	8.3	56.0	3.3	24.6	+	−
	GBN	A3-Bot	*1.1*	*5.8*	4.3	25.8	2.6	22.2	2.5	22.2	−	+
	Tali	A3-Bot	*1.8*	*6.3*	7.7	43.6	7.3	51.2	*1.4*	*12.5*	+	−
	Ratio1	A3-Bot	8.8	27.8	16.4	50.4	5.9	41.3	8.8	51.7	−	−
	Ratio2	A3-Bot	3.2	20.7	4.4	11.5	3.2	24.7	2.1	20.1	−	+
Ali-QTL4.1	PRO	A4-Mid	N	N	N	N	2.7	17.8	N	N	N	−
Ali-QTL5.1	Tali	A5-Mid	N	N	N	N	2.6	23.3	N	N	N	+
Ali-QTL5.2	Ratio2	A5-Mid	N	N	2.9	8.4	N	N	N	N	−	N
Ali-QTL6.1	NAP	A6-Mid	5.2	17.0	N	N	N	N	N	N	+	N
	Tali	A6-Mid	4.0	15.6	N	N	N	N	N	N	+	N
Ali-QTL7.1	NAP	A7-Mid	N	N	3.0	9.2	N	N	*1.6*	*8.9*	−	+
	Tali	A7-Mid	N	N	2.4	12.6	N	N	2.9	32.4	−	+
Ali-QTL8.1	PRO	A8-Bot	N	N	N	N	3.7	25.9	N	N	N	−
Ali-QTL9.1	NAP	A9-Mid	N	N	2.6	7.0	N	N	N	N	+	N
Ali-QTL9.2	GBN	A9-Mid	3.2	19.7	N	N	*1.4*	*10.6*	N	N	+	−
Ali-QTL9.3	Ratio1	A9-Bot	2.0	5.7	N	N	N	N	N	N	+	N
	Ratio2	A9-Bot	N	N	2.6	6.7	N	N	N	N	−	N
	NAP	A9-Bot	N	N	N	N	*1.3*	*5.4*	2.2	14.8	N	+
Ali-QTL10.1	PRO	A10-Mid	9.2	43.3	N	N	N	N	N	N	−	N
	NAP	A10-Mid	2.0	6.0	N	N	N	N	*1.4*	*8.6*	+	−
	Tali	A10-Mid	2.5	9.4	N	N	N	N	*1.9*	*17.4*	+	−
Ali-QTL10.2	Ratio1	A10-Bot	N	N	2.9	5.8	N	N	N	N	+	N
Ind-QTL2.1	TIND	A2-Mid	2.6	12.7	*1.0*	*6.2*	N	N	N	N	−	N
Ind-QTL3.1	TIND	A3-Bot	N	N	N	N	N	N	2.0	18.1		+
Ind-QTL8.1	TIND	A8-Bot	3.2	16.4	*1.5*	*9.9*	N	N	N	N	+	N
Aro-QTL4.1	NAS	A4-Top	N	N	2.1	16.2	N	N	N	N	−	N
Aro-QTL4.2	NAS	A4-Mid	N	N	N	N	N	N	2.8	24.7	N	−
Aro-QTL7.1	NAS	A7-Top	N	N	N	N	3.1	24.9	N	N	N	+

[a]Ali(aliphatic), Ind (indole) or Aro(aromatic), followed by linkage group and relative position from the top of that linkage group(Ali-QTL1.2 etc).
[b]Abbreviations used are explained in Table 1. N, no QTL present; LOD scores and exp% in italics are nonsignificant, but are recorded, since at similar positions, significant QTL were detected in other populations/seasons.
[c]+, yellow sarson increases the GS content; −, yellow sarson decreases the GS content.

Indole glucosinolates are produced from tryptophan, while aromatic glucosinolates are produced from phenylalanine. As described previously, some correlations were identified between NAS and the aliphatic glucosinolates NAP and GBN with some indole glucosinolates (Table 3). This could indicate possible cross-talk between different glucosinolate biosynthesis pathways.

For indolic glucosinolates, the sum of 4ME, 4OH, GBC and NEO concentrations was used in QTL mapping because all the individual indolic glucosinolates are correlated with each other and present in very low concentrations. In DH38, two QTL were detected on A2 and A8 that were only significant in autumn 2004, while in DH30 one QTL for total indole glucosinolates was identified on A3 (Ind-QTL3.1), explaining 18% of the variation in spring 2005. Ind-QTL3.1 co-localized with the aliphatic glucosinolates Ali-QTL3.2 based on the map position (Fig. 3).

The only aromatic glucosinolate detected in this study is 2-phenylethyl glucosinolate (NAS), which was also identified in all 60 *B. rapa* accessions screened for glucosinolate content (data not shown). Aro-QTL4.1 was detected in DH38 in spring 2005, while in DH 30 two other QTL for NAS (Aro-QTL4.2 and Aro-QTL7.1) were identified on A4 and A7 in spring 2005 and in autumn 2004, respectively, each
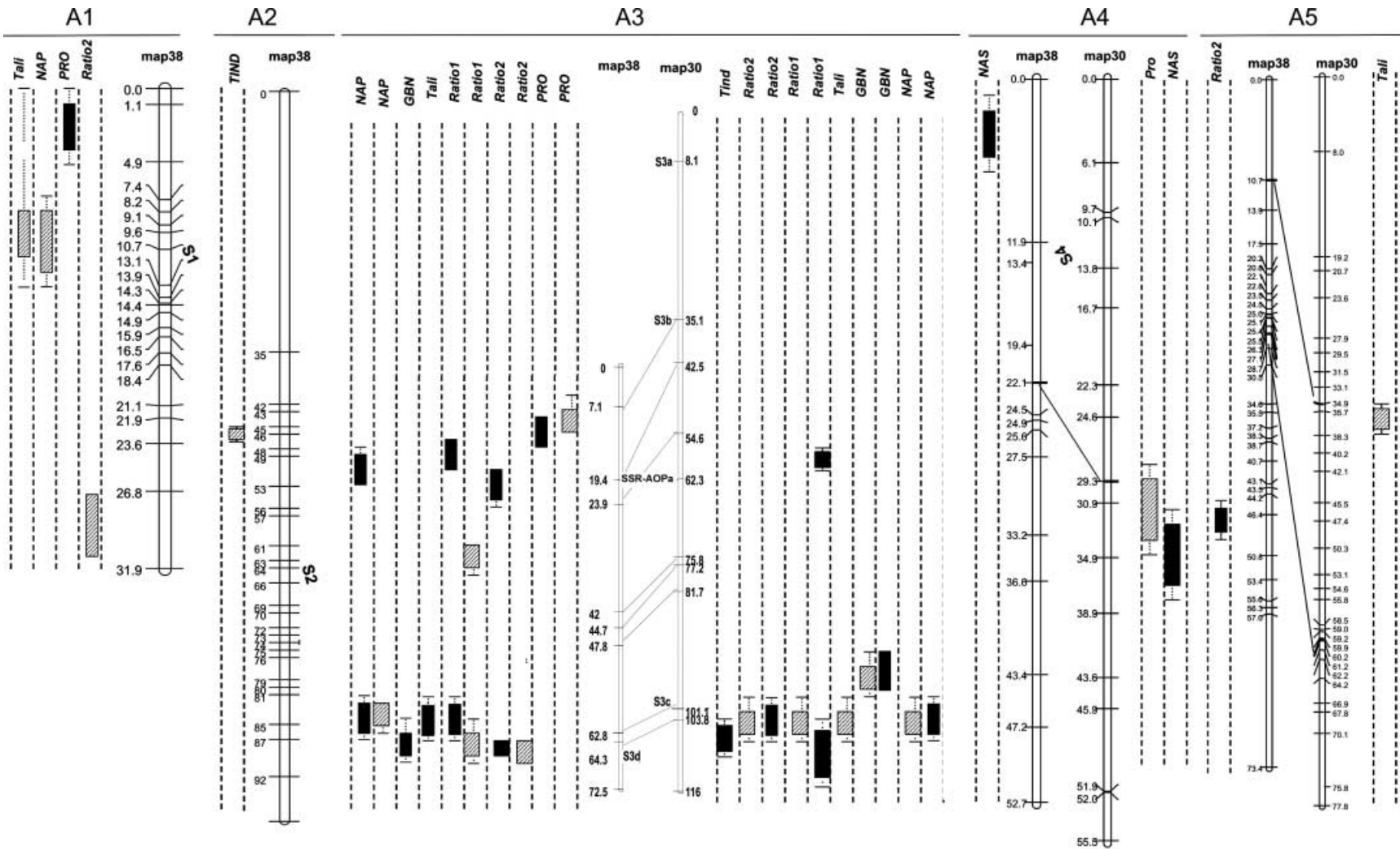
Fig. 3 Position of quantitative trait loci (QTL) of aliphatic, indolic and aromatic glucosinolate detected in the *Brassica rapa* double haploid (DH) populations DH38 and DH30. LOD profiles are graphically displayed using a map chart, including the LOD minus 1/LOD minus 2 support interval to approximate a 95% confidence interval. Gray bars, QTL for glucosinolates detected in autumn 2004; black bars, QTL for glucosinolates detected in spring 2005. A number of single sequence repeat (SSR) markers are indicated on the map: S1, BRMS096; S2, Na12H09; S3a, BRMS042; S3b, BRMS050; S3c, Na12E02; S3d, BRMS008; S4, Na10D09; S6, BRMS014; S7a, BRMS018; S7b, BRMS040; S7c,. Ra2A01; S7d, Ol12E03; S9, Ol10D08; S10, BrFLC1.
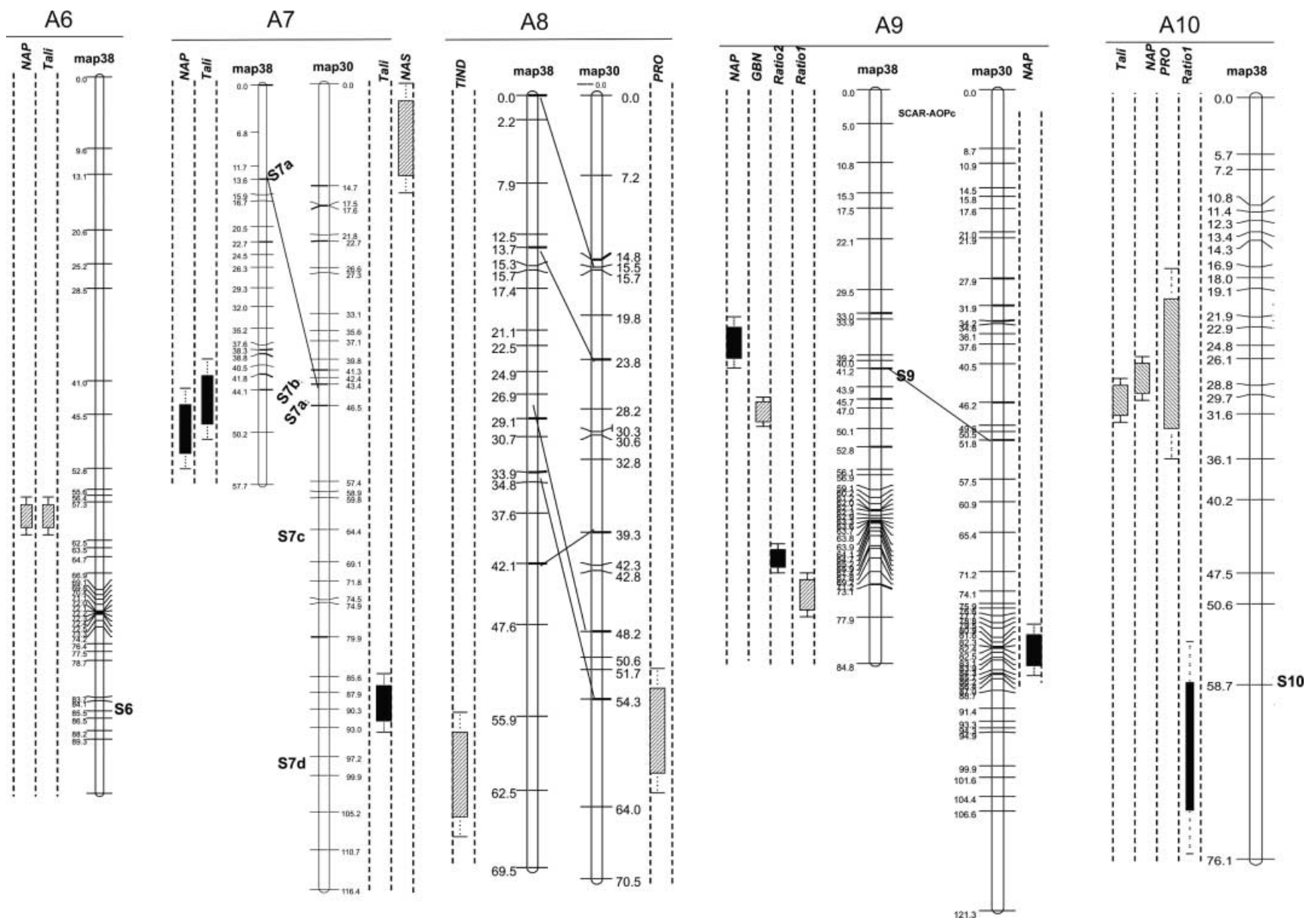
Fig. 3 continued

explaining 25% of the variation. Aro-QTL4.1 co-localized with the aliphatic glucosinolates Ali-QTL4.1 for PRO based on map alignment.

Discussion

Populations of DH lines have been used extensively for QTL mapping because they are immortal and completely homozygous. They have obvious advantages above F_2 populations, which are represented by single, often heterozygous, plants. However, recombination events between parental genomes are lower in DH lines compared with recombinant inbred lines (RILs). The latter are, however, hard to create and to maintain because of the self-incompatibility of most *B. rapa* accessions. We analyzed two DH populations which both shared yellow sarson R500 (YS-143) as male parent and had as female parents very different east Asian *B. rapa* types, providing ample variation for molecular and phenotypic traits, in contrast to many earlier DH populations that were made between either Chinese cabbage types or between oil types (summer oil , yellow sarson and winter oil types) (Wang *et al.*, 2004; Zhang *et al.*, 2005; Kim *et al.*, 2006;Suwabe *et al.*, 2006).

In order to compare our linkage map with public *B. rapa* maps, SSR primers were used to assign linkage groups to international standard A groups (www.brassica.info). One important criterion used during the selection of SSR loci was the requirement that each primer only has one map position in reference linkage maps. This avoids the ambiguity that specific SSR markers map at more, or different, positions in different *B. rapa* maps. The positions of SSR loci in the DH maps presented here were consistent with those in other published *Brassica* maps except for Na12A01, which mapped to A1 in DH38 and not to A4 (N4) as listed in the MBGP website (www.brassica.info). However Piquemal *et al.* (2005) showed that Na12A01 maps to at least three loci in the *B. napus* linkage map, one of them on N01, which corresponds to A1(Piquemal *et al.*, 2005). Marker coverage in our DH maps was good, with some clustering of markers on several linkage groups. The orientation of linkage groups A3 and A10 is different in published *B. rapa* genetic maps; one orientation in the F_2 linkage map of Kim *et al.* (2006) and the opposite orientation in the DH linkage map published by Choi *et al.* (2007). The orientation as published by Choi *et al.* (2007) is now the international standard, and is used in this paper. This orientation is similar to that of the *B. napus* maps used for comparative mapping published by Parkin *et al.* (2005). We need to state here that in our earlier publications, we also used DH-30 and DH-38 to map QTL for phytate, phosphate and developmental traits, and used the orientation as published by Kim *et al.* (2006), and thus A3 and A10 have different orientations compared with this paper (Lou *et al.*, 2007; Zhao *et al.*, 2008). Sixteen genetic loci controlled individual or total aliphatic glucosinolate accumulation and their ratios; four QTL are detected in both populations and four QTL are identified
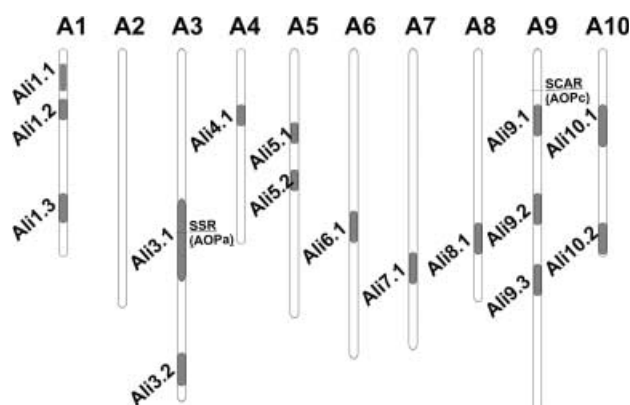

Fig. 4 Graphical presentation of aliphatic glucosinolate quantitative trait loci (QTL) positions in double haploid (DH) populations.

in both seasons, while others are population/season-specific. Not all QTL were identified in both experiments, which partly illustrates environmental effects on glucosinolate composition. Additional causes may be different developmental stages as a result of different growth rates in autumn and spring, and the fact that population size was relatively small with no replications within the experiments. Ali-QTL3.2 at the bottom of A3 is a major QTL detected in both populations and both seasons for the individual aliphatic glucosinolates and their ratios, which co-localizes with Ind-QTL3.2 for total indole glucosinolates detected in DH-30 spring 2005. This is the only clear co-localization for different classes of glucosinolates. Ali-QTL3.2 also affects the ratio between 4-C and 5-C aliphatic glucosinolates (ratio 1: (NAP + PRO)/TAli), explaining 28–50% of the variation in both populations and seasons, while AliQTL3.1 explains between 6 and 12% for ratio 1, also in both seasons and populations. AliQTL3.2 explained between 12 and 25% of the variation for ratio 2 between PRO and PRO + NAP in both populations and seasons, while AliQTL3.1 affects these ratios only in DH38, explaining 28% in 2005. These data suggest a role for enzymes like 2-oxoglutarate-dependent dioxygenases (AOP) (Kliebenstein *et al.*, 2001b), involved in side-chain modification.

The genome of *B. rapa* has triplicate counterparts to corresponding homeologous segments of Arabidopsis; as a result, many genes are present in duplicate or triplicate in the *B. rapa* genome, and multiple copies of genes may contribute to phenotypic diversity (Parkin *et al.*, 2005; Schranz *et al.*, 2006). Based on these data, we can predict the putative location of homologues of Arabidopsis genes involved in glucosinolate biosynthesis, such as *MAM* and *AOP*, but also transcription factors like Myb28 and Myb29 which regulate aliphatic glucosinolate biosynthesis (Hirai *et al.*, 2007) (Fig. 4). *AOP* is predicted to have two homeologous loci, one on A3 and another on A9 in the *B. rapa* genome, which was confirmed by the mapping of SSR-AOPa (an SSR physically linked to an AOP locus on BAC KBrB002P01) and SCAR-AOPc in the

DH populations on A3 and A9, respectively (Figs 1, 4). For *MAM* (segment Q) three paralogues are predicted on A2, A6 and A9, and for Myb29 (segment R) three paralogues on A2, A3 and A10, while Myb28 (segment X) paralogues are predicted on chromosomes 2 and 6. The QTL analysis reported in this paper is too imprecise to associate specific QTL with candidate gene loci. However, we can use the comparisons between our QTL and the map positions of predicted glucosinolate biosynthesis genes in reference maps based on SSR markers and map distance to suggest candidate genes. To obtain proof of the hypothesis, the candidate genes need to be mapped on our DH maps and allelic variation needs to be investigated.

In a recent study, *Myb28*, *Myb76* and *Myb29* were identified as positive regulators of aliphatic glucosinolate accumulation, and it was suggested that *Myb29* had a preference for short-chain aliphatic glucosinolates as identified in this study (Hirai *et al.*, 2007; Gigolashvili *et al.*, 2007c). Two other studies in a population of 403 Arabisopsis Bay-0 × Sha RILs also investigated the roles of the R2R3 Myb transcription factors *MYB28*, *MYB29* and *MYB76*, and also concluded that these genes evolved both overlapping and specific regulatory capacities. By genetic co-localization, *MYB28* was identified as candidate gene for both aliphatic glucosinolate QTL and expression QTL of genes involved in aliphatic glucosinolate biosynthesis. This suggests that these QTL may be explained by variation in the expression of this Myb gene (Sonderby *et al.*, 2007; Wentzell *et al.*, 2007). In *B. rapa* we do not see evidence for co-localization of *MYB28* or *MYB29* with the major aliphatic QTL AliQTL3.2. However, synthetic relations may suggest the orthologous gene of *MYB28* as a putative candidate gene for AliQTL6.1, explaining around 15% of the variation in NAP and Tali in DH 38 (both seasons), and an orthologue of *MYB29* as a putative candidate gene for AliQTL10.2, explaining 6% for ratio 1 (elongation) in DH38 (autumn 2004). The mapped *AOP* locus on A3 maps in the vicinity of Ali-QTL3.1 (ratio 1; elongation, ratio 2; modification, NAP, PRO) may thus account for these QTL. Another candidate for both AliQTL9.1 and Ali-QTL6.1 that needs to be tested by genetic mapping is *MAM*, since the Arabidopsis orthologue maps at a synthetic position. Genetic mapping of the *Brassica* paralogues is under way to confirm co-localization of candidate genes with QTL intervals.

In conclusion, many QTL for aliphatic glucosinolates are identified in two *B. rapa* DH populations, with the major Ali-QTL3.2 on A3 co-localizing with a QTL for indole glucosinolates. From 16 genetic loci controlling individual or total aliphatic glucosinolate accumulation and their ratios, four QTL are detected in both populations. Another four QTL are identified in both seasons while others are population/season-specific. Studies that combine gene expression profiles with metabolite profiles are presently being conducted in DH populations and these studies, together with the mapping of candidate genes, may provide more information on genes underlying the QTL. As the genetic basis of glucosinolate biosynthesis becomes clear, it will be feasible to breed for *Brassica* vegetables with different types and concentrations of glucosinolates by marker-assisted selection or by transgenic approaches.

References

Bak S, Feyereisen R. 2001. The involvement of two P450 enzymes, CYP83B1 and CYP83A1, in auxin homeostasis and glucosinolate biosynthesis. *Plant Physiology* **127**: 108–118.

Blau PA, Feeny P, Contardo L. 1978. Allylglucosinolate and herbivorous caterpillars: a contrast in toxicity and tolerance. *Science* **200**: 1296–1298.

Branca F, Li G, Goyal S, Quiros CF. 2002. Survey of aliphatic glucosinolates in sicilian wild and cultivated Brassicaceae. *Phytochemistry* **59**: 717–724.

Celenza JL, Quiel JA, Smolen GA, Merrikh H, Silvestro AR, Normanly J, Bender J. 2005. The arabidopsis ART1 MYB transcription factor controls indolic glucosinolate homeostasis. *Plant Physiology* **137**: 253–262.

Choi SR, Teakle GR, Plaha P, Kim JH, Allender CJ, Beynon E, Piao ZY, Soengas P, Han TH, King GJ *et al.* 2007. The reference genetic linkage map for the multinational Brassica rapa genome sequencing project. *Theoretical and Applied Genetics* **115**: 777–792.

Coventry J, Kott L, Beversdorf WD. 1988. *Manual for microspore culture technique for Brassica napus*. Guelph, Ontario, Canada: Department of Crop Science, University of Guelph.

Custers JBM, Cordewener JHG, Fiers MA, Maassen BTH, Van Lookeren Campagne MM, Liu CM. 2001. Androgenesis in Brassica; a model system to study the initiation of plant embryogenesis. In: Bhojwani SS, Soh WY, eds. *Current trends in the embryology of angiosperms*. Dordrecht, the Nertherlands: Kluwer Academic Publishers, 451–470.

Custers JBM, Cordewener JHG, Nöllen Y, Dons HJM, Van Lookeren Campagne MM. 1994. Temperature controls both gametophytic and sporophytic development in microspore cultures of *Brassica napus*. *Plant Cell Reports* **13**: 267–271.

Fahey JW, Stephenson KK, Talalay P. 1998. Glucosinolates, myrosinase, and isothiocyanates: three reasons for eating Brassica vegetables. *ACS Symposium Series* **701**: 16–22.

Fahey JW, Zalcmann AT, Talalay P. 2001. The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry* **56**: 5–51.

Field B, Cardon G, Traka M, Botterman J, Vancanneyt G, Mithen R. 2004. Glucosinolate and amino acid biosynthesis in Arabidopsis. *Plant Physiology* **135**: 828–839.

Field B, Furniss C, Wilkinson A, Mithen R. 2006. Expression of a Brassica isopropylmalate synthase gene in Arabidopsis perturbs both glucosinolate and amino acid metabolism. *Plant Molecular Biology* **60**: 717–727.

Gao M, Li G, Yang B, McCombie WR, Quiros CF. 2004. Comparative analysis of a Brassica BAC clone containing several major aliphatic

glucosinolate genes with its corresponding Arabidopsis sequence. *Genome* **47**: 666–679.

Gigolashvili T, Berger B, Mock HP, Muller C, Weisshaar B, Flugge UI. 2007a. The transcription factor HIG1/MYB51 regulates indolic glucosinolate biosynthesis in *Arabidopsis thaliana*. *Plant Journal* **50**: 886–901.

Gigolashvili T, Engqvist M, Yatusevich R, Muller C, Flugge UI. 2007c. HAG2/MYB76 and HAG3/MYB29 exert a specific and coordinated control on the regulation of aliphatic glucosinolate biosynthesis in *Arabidopsis thaliana*. *New Phytologist* **177**: 627–642.

Gigolashvili T, Yatusevich R, Berger B, Muller C, Flugge UI. 2007b. The R2R3-MYB transcription factor HIG1/MYB28 is a regulator of methionine-derived glucosinolate biosynthesis in *Arabidopsis thaliana*. *Plant Journal* **51**: 247–261.

Grubb CD, Zipp BJ, Ludwig-Muller J, Masuno MN, Molinski TF, Abel S. 2004. Arabidopsis glucosyltransferase UGT74B1 functions in glucosinolate biosynthesis and auxin homeostasis. *Plant Journal* **40**: 893–908.

Halkier BA, Du L. 1997. The biosynthesis of glucosinolates. *Trends in Plant Science* **2**: 425–431.

Halkier BA, Gershenzon J. 2006. Biology and biochemistry of glucosinolates. *Annual Review Plant Biology* **57**: 303–333.

He H, Fingerling G, Schnitzler WH. 2000. Glucosinolate contents and patterns in different organs of Chinese cabbages, Chinese kale (*Brassica alboglabra* bailey) and Choy sum (*Brassica campestris* L. ssp. *chinensis* var. *Utilis* tsen et lee). *Journal of Applied Botany* **74**: 21–25.

Heidel AJ, Clauss MJ, Kroymann J, Savolainen O, Mitchell-Olds T. 2006. Natural variation in MAM within and between populations of *Arabidopsis lyrata* determines glucosinolate phenotype *Genetics* **173**: 1629–1636.

Hirai MY, Klein M, Fujikawa Y, Yano M, Goodenowe DB, Yamazaki Y, Kanaya S, Nakamura Y, Kitayama M, Suzuki H *et al*. 2005. Elucidation of gene-to-gene and metabolite-to-gene networks in Arabidopsis by integration of metabolomics and transcriptomics. *Journal of Biological Chemistry* **280**: 25590–25595.

Hirai MY, Sugiyama K, Sawada Y, Tohge T, Obayashi T, Suzuki A, Araki R, Sakurai N, Suzuki H, Aoki K *et al*. 2007. Omics-based identification of Arabidopsis MYB transcription factors regulating aliphatic glucosinolate biosynthesis. *Proceedings of the National Academy of Sciences, USA* **104**: 6478–6483.

Howell PM, Sharpe AG, Lydiate DJ. 2003. Homoeologous loci control the accumulation of seed glucosinolates in oilseed rape (*Brassica napus*). *Genome* **46**: 454–460.

Keurentjes JJ, Fu J, de Vos CH, Lommen A, Hall RD, Bino RJ, van der Plas LH, Jansen RC, Vreugdenhil D, Koornneef M. 2006. The genetics of plant metabolism. *Nature Genetics* **38**: 842–849.

Kim JS, Chung TY, King GJ, Jin M, Yang T-J, Jin Y-M, Kim H-I, Park B-S. 2006. A sequence-tagged linkage map of *Brassica rapa*. *Genetics* **174**: 29–39.

Kliebenstein DJ, Gershenzon J, Mitchell-Olds T. 2001c. Comparative quantitative trait loci mapping of aliphatic, indolic and benzylic glucosinolate production in *Arabidopsis thaliana* leaves and seeds. *Genetics* **159**: 359–370.

Kliebenstein DJ, Kroymann J, Brown P, Figuth A, Pedersen D, Gershenzon J, Mitchell-Olds T. 2001a. Genetic control of natural variation in Arabidopsis glucosinolate accumulation. *Plant Physiology* **126**: 811–825.

Kliebenstein DJ, Lambrix VM, Reichelt M, Gershenzon J, Mitchell-Olds T. 2001b. Gene duplication in the diversification of secondary metabolism: tandem 2-oxoglutarate-dependent dioxygenases control glucosinolate biosynthesis in Arabidopsis. *Plant Cell* **13**: 681–693.

Kliebenstein DJ, Pedersen D, Barker B, Mitchell-Olds T. 2002. Comparative analysis of quantitative trait loci controlling glucosinolates, myrosinase and insect resistance in *Arabidopsis thaliana*. *Genetics* **161**: 325–332.

Koornneef M, Alonso-Blanco C, Vreugdenhil D. 2004. Naturally occurring genetic variation in *Arabidopsis thaliana*. *Annual Review Plant Biology* **55**: 141–172.

Kroymann J, Donnerhacke S, Schnabelrauch D, Mitchell-Olds T. 2003. Evolutionary dynamics of an Arabidopsis insect resistance quantitative trait locus. *Proceedings of the National Academy of Sciences, USA* **100**(Suppl. 2): 14 587–14 592.

Kroymann J, Textor S, Tokuhisa JG, Falk KL, Bartram S, Gershenzon J, Mitchell-Olds T. 2001. A gene controlling variation in Arabidopsis glucosinolate composition is part of the methionine chain elongation pathway. *Plant Physiology* **127**: 1077–1088.

Levy M, Wang Q, Kaspi R, Parrella MP, Abel S. 2005. Arabidopsis IQD1, a novel calmodulin-binding nuclear protein, stimulates glucosinolate accumulation and plant defense. *Plant Journal* **43**: 79–96.

Li G, Quiros CF. 2002. Genetic analysis, expression and molecular characterization of BoGSL-ELONG, a major gene involved in the aliphatic glucosinolate pathway of Brassica species. *Genetics* **162**: 1937–1943.

Li G, Quiros CF. 2003. In planta side-chain glucosinolate modification in Arabidopsis by introduction of dioxygenase Brassica homolog BoGSL-ALK. *Theoretical and Applied Genetics* **106**: 1116–1121.

Lou P, Zhao J, Kim JS, Shen S, Pino del Carpio D, Song X, Jin M, Vreugdenhil D, Wang X, Koornneef M *et al*. 2007. Quantitative trait loci for flowering time and morphological traits in multiple populations of *Brassica rapa*. *Journal of Experimental Botany* **58**: 4005–4016.

Mahmood T, Ekuere U, Yeh F, Good AG, Stringam GR. 2003. Molecular mapping of seed aliphatic glucosinolates in *Brassica juncea*. *Genome* **46**: 753–760.

Mikkelsen MD, Naur P, Halkier BA. 2004. Arabidopsis mutants in the *C-S* lyase of glucosinolate biosynthesis establish a critical role for indole-3-acetaldoxime in auxin homeostasis. *Plant Journal* **37**: 770–777.

Mithen R. 2001. Glucosinolates – biochemistry, genetics and biological activity. *Plant Growth Regulator* **34**: 91–103.

Mithen RF, Dekker M, Verkerk R, Rabot S, Johnson IT. 2000. The nutritional significance, biosynthesis and bioavailability of glucosinolates in human foods. *Journal of the Science of Food and Agriculture* **80**: 967–984.

Murray MG, Thompson WF. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* **8**: 4321–4325.

Neo JC, Rose P, Ong CN, Chung MC. 2005. Beta-phenylethyl isothiocyanate mediated apoptosis: a proteomic investigation of early apoptotic protein changes. *Proteomics* **5**: 1075–1082.

Padilla G, Cartea ME, Velasco P, de Haro A, Ordas A. 2007. Variation of glucosinolates in vegetable crops of *Brassica rapa*. *Phytochemistry* **68**: 536–545.

Parkin IA, Gulden SM, Sharpe AG, Lukens L, Trick M, Osborn TC, Lydiate DJ. 2005. Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. *Genetics* **171**: 765–781.

Pfalz M, Vogel H, Mitchell-Olds T, Kroymann J. 2007. Mapping of QTL for resistance against the crucifer specialist herbivore pieris brassicae in a new arabidopsis inbred line population, da(1)-12xei-2. *PLoS ONE* **2**: e578.

Piotrowski M, Schemenewitz A, Lopukhina A, Muller A, Janowitz T, Weiler EW, Oecking C. 2004. Desulfoglucosinolate sulfotransferases from *Arabidopsis thaliana* catalyze the final step in the biosynthesis of the glucosinolate core structure. *Journal of Biological Chemistry* **279**: 50717–50725.

Piquemal J, Cinquin E, Couton F, Rondeau C, Seignoret E, Doucet I, Perret D, Villeger MJ, Vincourt P, Blanchard P. 2005. Construction of an oilseed rape (*Brassica napus* L.) genetic map with SSR markers. *Theoretical and Applied Genetics* **111**: 1514–1523.

Ramchiary N, Bisht NC, Gupta V, Mukhopadhyay A, Arumugam N, Sodhi YS, Pental D, Pradhan AK. 2007. QTL analysis reveals context-dependent loci for seed glucosinolate trait in the oilseed *Brassica juncea*:

importance of recurrent selection backcross scheme for the identification of 'true' QTL. *Theoretical and Applied Genetics* **116**: 77–85.

Reichelt M, Brown PD., Schneider B, Oldham NJ, Stauber E, Tokuhisa J, Kliebenstein DJ, Mitchell-Olds T, Gershenzon J. 2002. Benzoic acid glucosinolate esters and other glucosinolates from *Arabidopsis thaliana*. *Phytochemistry* **59**: 663–671.

Rosa EAS, Hesney RK, Fenwick GR, Portas C. 1997. Glucosinolate in crop plant. *Horticulture Review* **19**: 99–215.

Schranz ME, Lysak MA, Mitchell-Olds T. 2006. The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. *Trends in Plant Science* **11**: 535–542.

Skirycz A, Reichelt M, Burow M, Birkemeyer C, Rolcik J, Kopka J, Zanor MI, Gershenzon J, Strnad M, Szopa J et al. 2006. DOF transcription factor AtDof1.1 (OBP2) is part of a regulatory network controlling glucosinolate biosynthesis in Arabidopsis. *Plant Journal* **47**: 10–24.

Sonderby IE, Hansen BG, Bjarnholt N, Ticconi C, Halkier BA, Kliebenstein DJ. 2007. A systems biology approach identifies a R2R3 MYB gene subfamily with distinct and overlapping functions in regulation of aliphatic glucosinolates. *PLoS ONE* **2**: e1322.

Stam P. 1993. Construction of integrated genetic linkage maps by means of a new computer package: joinmap. *Plant Journal* **3**: 739–744.

Suwabe K, Tsukazaki H, Iketani H, Hatakeyama K, Kondo M, Fujimura M, Nunome T, Fukuoka H, Hirai M, Matsumoto S. 2006. Simple sequence repeat-based comparative genomics between *Brassica rapa* and *Arabidopsis thaliana*: the genetic origin of clubroot resistance. *Genetics* **173**: 309–319.

Talalay P, Fahey JW. 2001. Phytochemicals from cruciferous plants protect against cancer by modulating carcinogen metabolism. *Journal of Nutrition* **131**: 3027S–3033S.

Van Ooijen JW. 2004. *Mapqtl5, software for mapping of quantitative trait loci in experimental populations*. Wageningen, the Netherlands: Kyazma BV.

Van Ooijen JW. 2006. *Joinmap4, software for calculation of genetic linkage maps in experimental populations*. Wageningen, the Netherlands: Kyazma B.V.

Van Os H, Stam P, Visser RG, Van Eck HJ. 2005. Record: a novel method for ordering loci on a genetic linkage map. *Theoretical and Applied Genetics* **112**: 30–40.

Verkerk R, Dekker M, Jongen WMF. 2001. Post-harvest increase of indolyl glucosinolates in response to chopping and storage of Brassica vegetables. *Journal of the Science of Food and Agriculture* **81**: 953–958.

Voorrips RE. 2002. Mapchart: software for the graphical presentation of linkage maps and QTLs. *Journal of Heredity* **93**: 77–78.

Vos P, Hogers R, Bleeker M, Reijans M, van der Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M et al. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* **23**: 4407–4414.

Wang M, Meng XD, Liu XC, Zhao XY, Fan ZC. 2004. A linkage map construction for chinese cabbage based on AFLP markers using DH population. *Acta Agricultural Boreali-Sinic* **19**: 28–33.

Wentzell AM, Rowe HC, Hansen BG, Ticconi C, Halkier BA, Kliebenstein DJ. 2007. Linking metabolic QTLs with network

and cis-eQTLs controlling biosynthetic pathways. *PLoS Genetic* **3**: 1687–1701.

Wittstock U, Halkier BA. 2000. Cytochrome P450 CYP79A2 from *Arabidopsis thaliana* L. Catalyzes the conversion of L-phenylalanine to Phenylacetaldoxime in the biosynthesis of benzylglucosinolate. *Journal of Biological Chemistry* **275**: 14 659–14 666.

Wittstock U, Halkier BA. 2002. Glucosinolate research in the Arabidopsis era. *Trends in Plant Science* **7**: 263–270.

Yang TJ, Kim JS, Kwon SJ, Lim KB, Choi BS, Kim JA, Jin M, Park JY, Lim MH, Kim HI et al. 2006. Sequence-level analysis of the diploidization process in the triplicated flowering locus C region of *Brassica rapa*. *Plant Cell* **18**: 1339–1347.

Zhang Y, Kensler TW, Cho CG, Posner GH, Talalay P. 1994. Anticarcinogenic activities of sulforaphane and structurally related synthetic norbornyl isothiocyanates. *Proceedings of the National Academy of Sciences, USA* **91**: 3147–3150.

Zhang XF, Wang XW, Lou P, Zhang XW, Wang YQ, Yuan YX, Zhao JJ, Sun RF. 2005. An AFLP-based genetic linkage map of chinese cabbage using double haploid(DH) population. *Acta Horticulturea Sinica* **32**: 443–448.

Zhao J, Jamar D, Lou P, Wang Y, Wu J, Wang X, Bonnema G, Koornneef M, Vreugdenhil D. 2008. QTL analysis of phytate and phosphate concentrations in seeds and leaves of *Brassica rapa*. *Plant, Cell & Environment* doi: 10.1111/j.1365-3040.2008.01791.x

Zhao J, Paulo MJ, Diaan J, Lou P, van Eeuwijk F, Bonnema G, Vreugdenhil D, Koornneef M. 2007. Association mapping of leaf traits, flowering time, and phytate content in *Brassica rapa*. *Genome* **50**: 963–973.

Zhao J, Wang X, Deng B, Lou P, Wu J, Sun R, Xu Z, Vromans J, Koornneef M, Bonnema G. 2005. Genetic relationships within *Brassica rapa* as inferred from AFLP fingerprints. *Theoretical and Applied Genetics* **110**: 1301–1314.

Supporting Information

Additional supporting information may be found in the online version of this article.

Table S1 Correlation coeffecient of all the individual glucosinolates in two experiments

Table S2 Detail information of quantitative trait loci (QTL) detected in DH38 and DH30

Please note: Blackwell Publishing are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.